The retina: a fascinating object of study for a physicist

Bruno Cessac

The retina: a fascinating object of study for a physicist

Bruno Cessac

Abstract I briefly present joint research where ideas and methods from theoretical physics can be applied to better understand the behaviour of the retina in normal, developmental and pharmacologically controlled conditions.

1 Introduction

Our visual system has astonishing capacities, from the rapid extraction of the main features of a visual scene, to higher level tasks like reading or face recognition. Our vision starts from the retina. This tiny membrane, only a few hundred microns thick, covering 75% of the internal ocular globe performs fundamental yet complex tasks. Although its primary function is to convert the photons from the outer world into sequences of action potentials (spike trains), encoding the visual scene, and conveying them to the visual cortex where they will be "decoded", the retina is not a mere camera. In recent years, researchers have indeed discovered that the retina "is smarter than scientists believed" [1].

In this paper, I would like to share with the reader the fascination of the retina for a physicist, working for years in the field of dynamical systems theory and statistical physics applied to "complex systems", especially neuronal models. Working with biologists and retinal specialists, I have discovered a beautiful object of studies both from the applied and theoretical

Bruno Cessac
Université Côte d'Azur, Biovision research project, Inria, 2004 route des Lucioles, 06902 Sophia-Antipolis e-mail: bruno.cessac@inria.fr

physics point of view. The retinal machinery (neurons, synapses, ion transport, light conversion from photoreceptors, etc.) is governed by physics. Yet the extrapolation of physical methods from theoretical physics (mean-field methods, transport equations, Gibbs distributions, etc.) raises several interesting questions that I have been confronted with during my research, and which I want to briefly present in this paper.

2 The retina structure

Throughout this paper, I will use "computer-oriented" language to deal with the retina: information, circuits, code, decode, computation, etc. This is a contemporary view, largely influenced by our computer-based society. Although this analogy is useful - it eases explanations and provides fruitful paradigms – it has its limitations, which are stressed throughout this paper.

The retina, much like the rest of our body, especially the brain, has an evident problem: it can't tolerate large variations in temperature. Especially, the Joule effect has to be strongly limited. As a consequence, neurons, which are cells producing electric currents, do not use electrons, instead they use ion transfer (sodium, potassium, calcium, chloride, etc.). The currents produced this way are small (of order $1-100$ pA), as well as voltage variations (~ 100 mV), thus with an electric power of order pW. Even if there are many neurons of different types in the retina (of order 10^8, including photoreceptors), the total heat production is quite small compared to a computer that would perform the same tasks. However, ions are quite slow, and the corollary is that the retina has to use massive parallel computations to perform complex tasks in a short time. This computation is achieved via neurons, but also by synapses: the synaptic organization of the retina plays a central role in its abilities (e.g., http://webvision.med.utah.edu/book/part-i-foundations/simple-anatomy-of-the-retina/).

The retina converts photons into variations of electric potential (phototransduction) via **photoreceptor** cells: rods (about 130 million) ensure eyesight in poor illumination; cones (about 7 million) ensure central vision and colour perception. Phototransduction is a very efficient mechanism as a single photon can produce a visual effect. This is due to a complex cascade of molecular mechanisms with a huge multiplicative effect. At the other side of the retina one finds **retinal ganglion cells** (RGCs), the final stage of retina encoding, as these are the cells that emit action potentials (spikes) via their axons (which constitute the optic nerve) to the visual cortex via the thalamus. There are about 1.2 to 1.5 million RGCs in the human retina. On average, each RGC receives inputs from about 100 rods and cones. These numbers

vary greatly among individuals and as a function of retinal location. In between, one finds 3 cells types: horizontal, bipolar and amacrine cells. Unlike most neurons, these cells communicate via graded potentials, rather than action potentials. **Horizontal cells** are laterally interconnecting neurons, helping integrate and regulate the input from multiple photoreceptors. **Bipolar** cells transmit the signals from the photoreceptors or the horizontal cells, and pass them on to the ganglion cells directly or indirectly (via **amacrine cells**).

The retina has therefore both a feed-forward structure (from photoreceptors to ganglion cells) and a lateral structure (due to horizontal and amacrine cells). This generates different types of neural circuits, which enable the RGCs to efficiently process local visual information such as dim light, small responses to single photon absorption, segregating moving objects, filtering the movement of body, head, or eye, motion extrapolation, detection of approaching motion, surprise at the missing element in the sequence. Many of these computations match the evident challenge of animals: to detect moving objects and locate them correctly; to struggle with a constantly moving image sensor; and to predict the future and adapt to changing conditions [1]. Thus, the thalamus and visual cortex receive not a computer-like pixel representation of the image, but a set of features processed via nonlinear mechanisms that researchers try to identify [1].

The optic nerve is therefore like an optical fiber with several millions of channels - the axon of each RGC - conveying a local spatio-temporal information encoded by sequences of spikes, decoded by the brain. However, in contrast to computers, the code has variability and, nevertheless, robustness. First, several presentations of the same visual stimulus do not trigger the same sequence of spikes although some statistical regularity is observed (typically, a given RGC type fires more intensively when a specific stimulus is presented). Second, there is not a unique coding strategy. RGCs convey part of the information independently from each other through their firing rates, or timing of spikes. But they share information because the spatial regions that they scan have overlaps (a photoreceptor contributes the activity of several RGCs) inducing stimuli-induced correlations in their response. In addition, the lateral connections from horizontal and amacrine cells induce indirect interactions between RGCs. Therefore, RGCs presumably also encode information at a population level. This "population coding" presents several advantages: redundancy, reduction of uncertainty, simultaneous coding of different stimulus attributes, fast response, etc. It is however a contemporary challenge to understand it.

3 Population coding and statistical physics

Current acquisition technologies (Multi-Electrodes Array, MEA) allow us to simultaneously record several thousands of RGCs in response to a visual scene, providing a contemporary challenge: to try and decipher the visual scene from the RGC spikes and thereby infer coding strategies of the visual system. Part of this information can be recovered by assuming that cells encode information independently. This allows one to design "decoders" based on firing rate, spike latency, rank order, etc. Yet, the decoders built this way have many fitting parameters and their efficiency may vary with the visual stimulus. In addition, it has been shown [2] that a part of the information is carried by the (weak) correlations between RGCs suggesting that population coding takes place.

For a modeller, it seems clear that a population of connected neurons submitted to an external stimulus will produce a correlated response at the population level. We have made a mathematical analysis of this aspect in [3, 4], using Integrate and Fire models. We have shown that the population statistics are described by a variable length Markov chain where transition probabilities can be explicitly written: they depend on neuron connectivity, on the stimulus and on spike history in a similar fashion as the so-called Generalized Linear Models [5].

Such Markov chains are closely related to what physicists call "Gibbs distributions", initially introduced by Boltzmann and Gibbs to establish a link between microscopic dynamics of particles and thermodynamics. Gibbs distributions are probabilities of exponential form where the term in the exponential is, in physics, proportional to the energy; the form of the energy is constrained by the forces involved in the problem and defines a statistical model or an "ensemble". More generally, in the correspondence with Markov chain, the term in the exponential has not the interpretation of a physical energy, but we will call it "energy" as well, for simplicity. When dynamics are time-translation invariant ("stationarity"), Gibbs distributions are obtained by maximizing the statistical entropy under the constraint that the average of the observables defining the energy is fixed (Maximum Entropy Principle, MEP), but their definition via the equivalence with variable length Markov chains allows for non-stationary situations.

Using Gibbs distributions to analyze retina data and population coding has shown great success within the last decade. In particular, several important results have been obtained by using the MEP for an energy having the form of an Ising model, i.e., taking into account instantaneous pairwise interactions between neurons [2]. Extensions to more general energy forms have been considered too (triplet interactions [6], time delayed interactions [7]). In particular, we have developed efficient algorithms and soft-

ware, PRANAS, [8] allowing us to fit the parameters of a Gibbs distribution (whose energy form is given) from MEA data.

This "Gibbs" approach is appealing for a physicist. It would allow us to apply the powerful techniques and concepts from statistical physics to the analysis of the neural code. In addition, showing a canonical form of energy fitting well with retina data could be a step towards the "thermodynamics" of the retina: to explain the dynamics of a large population of RGCs by combining a few canonical observables. However, this approach raises several deep questions.

- **Which energy form?** In contrast to statistical physics/thermodynamics, the energy form for the retina cannot be inferred from first principles, so researchers are reduced to guess the form. Unfortunately, a mathematical analysis based on the mapping between Markov chains describing the neuronal dynamics and Gibbs distribution shows that the corresponding energy generically has a plethora of highly redundant observables [4]. We have proposed a method to eliminate these redundant terms from data analysis using information geometry, and we have shown experimentally that the degree of redundancy depends on the visual stimulus correlations [9].
- **Non stationarity.** Most Gibbs approaches, based on MEP, use the assumption of stationarity. To the contrary, the retina mainly responds to changes in a visual scene, i.e., transient, non-stationary stimuli. The MEP does not extend to this case. We have developed an approach, based on linear response theory, where a time-dependent stimulus is viewed as a perturbation of a stationary state (spontaneous activity) that can be characterized from data using MEP. In this case, the response to the stimulus can be written in terms of correlations of the stationary case (this an extension of the fluctuation-dissipation theorem of physics) [10].
- **Decoding.** Assume we are able to characterize the population statistics of RGC with a Gibbs distribution, how can we use it to decode the visual stimulus? Although some promising approaches have been proposed, this question seems far from being solved.

4 Retinal waves, retinal development and non linear dynamics

Immediately after birth the visual system of vertebrates is not yet effective. A complex, transient sequence of dynamical processes takes place starting a few days before birth, progressively enabling "the eyes to see" and stopping when vision is functional. A large part of this processing is due to waves of

electric activity ("retinal waves") spreading through the retina with a characteristic periodicity. This macroscopic phenomenon (i.e., occurring at the scale of the whole retina) originates from microscopic processes starting at the molecular level (ionic channels), inducing bursts of activity in specific cells, and spreading through the retina by virtue of cell connectivity. Retinal waves are classified into 3 consecutive stages, each having a specific role in visual system development. The transition between stages results from genetically programmed morphological changes. However, a part of this spatio-temporal activity and its transformation during development can be explained by generic mechanisms in nonlinear dynamics, as we describe here, focusing on stage II. This section is a summary of D. Karvouniari's thesis [11, 12], work conducted in collaboration with Institut de la Vision and InPhyNi (L. Gil).

Stage II retinal waves are due to spontaneous and periodic bursts of activity of specific retinal cells, the starburst amacrine cells (SACs), coupled by the excitatory neurotransmitter, acetylcholine (Ach). When a SAC is active (bursting) it releases acetylcholine; this can trigger the activation of post-synaptic cells. The membrane potential V of SAC i can be modeled as:

$$C\frac{dV_i}{dt} = I_{ion}(V_i, \bullet) + I_{sAHP}(V_i, R_i) + I_{Ach}(V_i, A_j), \tag{1}$$

where C is the membrane capacitance. The term $I_{ion}(V, \bullet)$ represents the sum of ionic currents involved in SACs bursting (mainly calcium and potassium), and depending on additional dynamical variables represented, for simplicity by the symbol $\bullet$ (see [11] for details); $I_{sAHP}(V_i, R_i)$ is a slow hyperpolarization potassium current, depending on a refractory variable R_i controlled by a cascade of kinetic processes; finally, $I_{Ach}(V, A_j)$ is the sum of excitatory acetylcholine currents due to active pre-synaptic cells j connected to i.

A bifurcation analysis of the model (1) shows that SACs can switch, by a saddle-node bifurcation, from a rest state to fast oscillations in the order of milliseconds (bursting). This arises when the current $I_{tot} = I_{sAHP} + I_{Ach}$ crosses from below a threshold value I_{SN}, depending on biophysical parameters (conductances, reversal potentials, etc.). Reciprocally, when the cell is bursting, it can go back to a rest state, by a homoclinic bifurcation, if I_{tot} crosses from above a threshold value I_{H_c}. In general, $I_{H_c} < I_{SN}$ but they differ by a few pA, so, for simplicity, we identify them from now. Thus, in short:

$$\text{if} \quad I_{tot} = I_{sAHP} + I_{Ach}, \begin{cases} < \theta, \text{SAC is at rest,} \\ > \theta, \text{SAC is active.} \end{cases} \tag{2}$$

The transition from rest to active is due to excitation from pre-synaptic active cells, via the excitatory current $I_{Ach}(V, A_j)$. The transition from active to

rest is due to the slow hyperpolarization current $I_{sAHP}(V_i, R_i)$ (having a negative sign). Indeed, when the cell is active, a complex mechanism involving calcium takes place, $I_{sAHP}(V_i, R_i)$ becomes more and more negative, leading the cell, after a few seconds and via a bifurcation, to a hyperpolarized rest state where it can not be excited for a long period (in the order of one minute), independently of the excitatory current I_{Ach} provided by the other cells.

Thus, wave propagation is due to a transition from rest to active state of SACs transmitted via Ach interactions. Waves are stopped by hyperpolarized regions corresponding to cells that have burst in a former wave. Therefore, each wave has to propagate into a landscape, imprinted by previous waves, with refractory regions and excitable regions. This landscape evolves slowly in time, on time scales that are longer than the SAC refractory period. This generates a spatial anisotropy where some cells are more active ("leaders") and some others more refractory. In this way, the mere dynamics generate a huge spatio-temporal variability, even if the cells are initially identical. This (biologically observed) variability is purely dynamical and does need to add extra mechanisms to be explained.

A non-linear wave propagation equation can be obtained, upon several approximations, considering SACs are located on a d-dimensional regular lattice, with spacing a, and nearest-neighbours interactions. The Ach conductance, Γ, considered now as a field in a d-dimensional continuum, obeys:

$$\frac{\partial \Gamma}{\partial t} = -\mu\,\Gamma + 2d\Omega H\left[\Gamma - \Gamma_c(R)\right] + a^2\Omega\Delta H\left[\Gamma - \Gamma_c(R)\right], \tag{3}$$

where μ is the Ach degradation rate, Ω the Ach production rate; Δ is the Laplacian operator; H is the Heaviside function, mimicking the threshold effect (2), and $\Gamma_c(R)$ is the critical threshold, derived from the bifurcation condition (2) in a refractory landscape characterized by the variable (field) R and depending upon the network history. This is a singular equation because one applies a Laplacian to a Heaviside function. It is however possible to smooth the Heaviside function to eliminate this singularity. This equation can be solved for simple refractory landscapes, but the general situation where R is a random landscape imprinted by waves history, is still under investigation.

The process of wave generation and propagation bares some similarity to forest fires introduced in the context of Self-Organized Criticality (SOC). SOC systems have the ability to self-organize into a state where characteristic events (avalanches) have power law distributions. This has lead some researchers to hypothesize that the retinal wave distribution (size or duration) could follow a power law [13]. Experimental evidence is not convincing though and would deserve a more elaborate analysis. Interestingly, eq.

(3) corresponds to a continuum limit of a SOC model (a sandpile) if the term $-\mu\Gamma + 2d\Omega H[\Gamma - \Gamma_c(R)] = 0$. This is very specific and non-generic situation. As a consequence, in our model, wave distributions are exponentials, except at a specific curve in the parameter space, given by this specific relation and related to the bifurcation condition (2); there the distribution is a power law. This suggests that SACs do not organize in a critical state unless some additional mechanism is added (like homeostasy), driving them towards the critical curve.

This model provides an example where one can construct the path from the molecular scale to the neuronal scale, to the macroscopic scale. The mathematical analysis allows us to explain bursting of SACs and wave propagation with simple mechanisms in nonlinear dynamics. In addition, it allows us to explicitly compute several important quantities, such as the wave speed.

But the main interest is the closeness to experiments. The model not only reproduces experimental facts, it also leads us to experimental predictions, some of which are on the way to be experimentally confirmed (see [11,12] on the role of kV3 channels on bursting of SACs during stage II). In particular, we are able to characterize how retinal wave structure is evolving during development when synaptic connections are modified. Likewise, the model is accurate enough to explain pharmacological manipulations (e.g., channels or synaptic terminal blocking).

5 Conclusion and perspectives

In this paper, I have given examples of research where concepts and methods from theoretical physics are used to understand retinal dynamics and how it encodes information. I would like now to briefly present further ongoing developments.

Amacrine cells and motion processing. When an object moves across the visual field our visual system is able to interpolate its trajectory and to filter much spurious information: eye-head-body movements, motion of the background. In particular, anticipation is absolutely essential to compensate the time lag of $30 - 100$ ms between the reception of photons in the retina and the response of the visual cortex. Part of the anticipation process starts in the retina and is explained by the non-linear response (gain control mechanism) of bipolar and ganglion cells [14]. This does not take into account the lateral connectivity of amacrine cells, which play an important role in

motion processing (differential motion, approaching motion, etc.). We want to understand the possible role of amacrine cells lateral connectivity in the retina in processing complex motion. In particular, we are seeking a specific transient signature in RGC correlations, in response to a moving object. We believe that the correlated response could provide more efficient processing of the object motion, especially trajectory anticipation and interpolation. We want to validate this hypothesis at the modelling level (PhD thesis of Selma Souihel) and experimental level, in collaboration with the Institut de la Vision and Institut des Neurosciences de la Timone, in the context of the Trajectory ANR.

Effect of pharmacologically switching a population of RGCs. At present, over 30 RGC sub-types have been identified, typically on the basis of common anatomical features or basic functions (e.g., sensitivity to motion, orientation, motion direction etc.). In collaboration with the University of Newcastle, in the context of E. Kartsaki's thesis, we want to investigate how different groups of RGCs contribute to the encoding of visual scenes. The project uses a pharmacogenetics approach (combined with MEA physiology, anatomy, computational modelling and behaviour) to reversibly silence subgroups of RGCs sharing gene expression through specific drugs (DREADD) activation. Removing an entire functional RGC group from the population response will shed light on the role these same cells play in population encoding of complex visual scenes and identify which information is lost, locally and globally.

References

1. T. Gollisch and M. Meister. *Neuron*, 65(2):150–164, January 2010.
2. E. Schneidman, M.J. Berry, R. Segev, and W. Bialek. *Nature*, 440(7087):1007–1012, 2006.
3. R. Cofré and B. Cessac. *Chaos, Solitons & Fractals*, 50(13):3, 2013.
4. R. Cofré and B. Cessac. *Phys. Rev. E*, 89(052117), 2014.
5. J. W. Pillow, Y. Ahmadian, and L. Paninski. *Neural Comput*, 23(1):1–45, 2011.
6. E. Ganmor, R. Segev, and E. Schneidman. *The journal of neuroscience*, 31(8):3044–3054, 2011.
7. J.C. Vasquez, O. Marre, A. Palacios, M.J Berry, and B. Cessac. *J. Physiol. Paris*, 106(3-4):120–127, May 2012.
8. B. Cessac, P. Kornprobst, S. Kraria, H. Nasser, D. Pamplona, G. Portelli, and T. Viéville. *Frontiers in Neuroinformatics*, 11:49, 2017.
9. R. Herzog, M.J. Escobar, R. Cofré, B. Cessac, and A. Palacios. networks. *PNAS (submitted)*, 2018.
10. B. Cessac and R. Cofré. submitted to Journal of Mathematical Biology, 2018.
11. D. Karvouniari, L. Gil, E. Orendorff, O. Marre, S. Picaud, and B. Cessac. *Scientific Reports, submitted*, 2018.

12. D. Karvouniari. *Retinal waves: theory, numerics, experiments*. PhD thesis, EDSTIC, 2018.
13. M.H. Hennig, C. Adams, D. Willshaw, and E. Sernagor. *J Neurosci*, 2009.
14. M.J. Berry, I.H. Brivanlou, T.A. Jordan, and M. Meister. *Nature*, 398(6725):334—338, 1999.